

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2005, 17:57:15 ; Search time 176 Seconds  
(without alignments)  
2155.970 Million cell updates/sec

Title: US-09-684-890A-2

Perfect score: 3887

Sequence: 1 MAAAGPVGDELWQTLFNP.....GEHIREMVKQINDIRHNVNF 741

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3817	98.2	741	1 NU88 HUMAN	Q99567 homo sapien
2	3461.5	89.1	742	1 NU88 RAT	O08658 rattus norv
3	3452	88.8	753	1 NU88 MOUSE	Q8ce00 mus musculu
4	2551.5	65.6	549	2 Q8BQF0	Q8bqf0 mus musculu
5	2417	62.2	728	2 Q707N0	Q707n0 xenopus lae
6	2397	61.7	726	2 Q6DDV7	Q6ddv7 xenopus lae
7	2386	61.4	726	2 Q707M9	Q707m9 xenopus lae
8	2133.5	54.9	462	2 Q8CAD9	Q8cad9 mus musculu
9	918	23.6	190	2 Q6Q7J4	Q6q7j4 sus scrofa
10	701	18.0	633	2 Q7PX24	Q7px24 anopheles g
11	646.5	16.6	702	1 NU88 DROME	Q9gyu8 drosophila
12	559.5	14.4	130	2 Q29177	Q29177 sus scrofa
13	318	8.2	342	2 Q7PKI7	Q7pk17 anopheles g
14	248.5	6.4	810	2 Q7PFR6	Q7pfr6 arabidopsis
15	211	5.4	890	2 Q7XPF0	Q7xpf0 oryza sativ
16	184.5	4.7	803	2 Q9P382	Q9p382 schizosacch
17	175.5	4.5	888	2 Q7RLX7	Q7rlx7 neurospora
18	166	4.3	980	2 Q9P0K7	Q9p0k7 homo sapien
19	166	4.3	983	2 Q7Z514	Q7z514 homo sapien
20	163	4.2	787	2 Q8NAB1	Q8nab1 homo sapien
21	163	4.2	980	2 Q7Z733	Q7z733 homo sapien
22	163	4.2	989	2 Q9P212	Q9p212 homo sapien
23	161	4.1	1253	2 Q86S00	Q86s00 homo sapien
24	158	4.1	344	2 Q9Y3T5	Q9y3t5 homo sapien
25	157	4.0	972	2 Q6V1W9	Q6v1w9 homo sapien
26	156	4.0	442	2 Q6AY97	Q6ay97 rattus norv
27	155	4.0	609	2 Q8TXA4	Q8txa4 methanopyru
28	154	4.0	964	2 Q8IIG7	Q8iig7 plasmodium
29	153	3.9	1366	2 Q6C134	Q6c134 varrowia li
30	152	3.9	920	2 Q7FAY1	Q7fay1 oryza sativ
31	149.5	3.8	442	2 Q9D8L5	Q9d8l5 mus musculu

32 149.5 3.8 894 2 Q7PAD5  
33 149.5 3.8 2954 2 Q42263  
34 147.5 3.8 978 1 RA50 AQUAE  
35 146 3.8 786 2 Q6BGH4  
36 145.5 3.7 1206 2 Q9NTC1  
37 144.5 3.7 1249 2 Q8KIN2  
38 143.5 3.7 793 2 Q8BKV3  
39 143.5 3.7 898 2 Q9Z220  
40 142.5 3.7 2663 1 CENE HUMAN  
41 141.5 3.6 446 2 Q66H66  
42 141.5 3.6 1302 2 Q80Y16  
43 140.5 3.6 713 1 NU82 YEAST  
44 140 3.6 876 1 RA50 METKA  
45 140 3.6 1107 2 Q7RME7

## ALIGNMENTS

### RESULT 1

ID NU88 HUMAN STANDARD; PRT; 741 AA.  
AC Q99567; Q9BWE5;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Nucleolar pore complex protein Nup88 (Nucleoporin Nup88) (88 kDa nuclear pore complex protein)  
DE Name=NUP88;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND INTERACTION WITH NUP214/CAN.  
RC TISSUE=Placenta;  
RX MEDLINE=97201523; PubMed=9049309; DOI=10.1093/emboj/16.4.807;  
RA Fornerod M., van Deursen J.M., van Baal S., Reynolds A., Davis D., Murti K.G., Franssen J., Grosveld G.;  
RT "The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and the novel nuclear pore component Nup88.";  
RL EMBO J. 16:807-816 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straubeberg R.L., Reingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McSwain K.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
RL Proc Natl. Acad. Sci U.S.A. 99:16899-16903 (2002).  
CC -!- FUNCTION: Essential component of nuclear pore complex.  
CC -!- SUBUNIT: Interacts with NUP214/CAN.  
CC -!- SUBCELLULAR LOCATION: Nuclear pore complex.  
CC -!- TISSUE SPECIFICITY: Ubiquitous.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: Y08612; CAA69904.1; -;  
 DR EMBL: BC000335; AAH00335.1; -;  
 DR Genew: HGNC:8067; NUP88.  
 DR MIM: 602552; -;  
 DR GO: GO:0005643; C:nuclear pore; TAS.  
 DR GO: GO:0005215; P:transporter activity; TAS.  
 KW Coiled coil; Nuclear protein; Protein transport; Transport.  
 FT DOMAIN 585 651 Coiled coil (potential).  
 FT CONFLICT 247 247 A -> D (in Ref. 1).  
 FT CONFLICT 291 301 GKLLGELPMPHP -> KWAVGSINAHAS (in Ref. 1).  
 FT CONFLICT 456 456 K -> R (in Ref. 1).  
 FT CONFLICT 518 518 P -> S (in Ref. 1).  
 SQ SEQUENCE 741 AA; 83541 MW; 954A8E2E203BC20B CRC64;

Query Match 98.2%; Score 3817; DB 1; Length 741;  
 Best Local Similarity 98.4%; Pred. No. 1.3e-205;  
 Matches 729; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAAAGPVGDEGELWQTLNPNHVVFLRLREGLNQSPTEAEKPASSSLPSSPPQLLTRNV 60  
 DB 1 MAAAGPVGDEGELWQTLNPNHVVFLRLREGLNQSPTEAEKPASSSLPSSPPQLLTRNV 60  
 QY 61 VFLGGLFLWDGEDSSFLVRLRGSGGGEPPALSOYORLLCINPPLFEIYQVLLSP 120  
 DB 61 VFLGGLFLWDGEDSSFLVRLRGSGGGEPPALSOYORLLCINPPLFEIYQVLLSP 120  
 QY 121 HVVALIGIGLWLELPKRWGNSEPEGGKSTVNCSTTPVAERFFTSSTLTUKHAAWP 180  
 DB 121 HVVALIGIGLWLELPKRWGNSEPEGGKSTVNCSTTPVAERFFTSSTLTUKHAAWP 180  
 QY 181 SEILDPHVLLTSDNVIIRIYSLREPOTPTNVIILSEAEESLVNKGRTAYTASLGATA 240  
 DB 181 SEILDPHVLLTSDNVIIRIYSLREPOTPTNVIILSEAEESLVNKGRTAYTASLGATA 240  
 QY 241 FDFGGLDVPKTLFGQNGKDEVVAYPLYLYENGETFLYISLLHSPGNWKAAGVIAHA 300  
 DB 241 FDFGGLDVPKTLFGQNGKDEVVAYPLYLYENGETFLYISLLHSPGNWKAAGVIAHA 300  
 QY 301 SAAEDNNGYDACAVALCLPCVPIILVIATBSGMLYHCVCVLEGEEDDHTSEKSWDSRIDLI 360  
 DB 301 SAAEDNNGYDACAVALCLPCVPIILVIATBSGMLYHCVCVLEGEEDDHTSEKSWDSRIDLI 360  
 QY 361 PSLYVFECEVELELALKASGEDDPFSDPSCPVKLHRDPKCPSRHYCTHEAGVHSVGLTW 420  
 DB 361 PSLYVFECEVELELALKASGEDDPFSDPSCPVKLHRDPKCPSRHYCTHEAGVHSVGLTW 420  
 QY 421 IHKLHFLGSDDEKDSLOELSTEQKCFVEHILCTPLCPROPAPIRGFWIVPDILGPTM 480  
 DB 421 IHKLHFLGSDDEKDSLOELSTEQKCFVEHILCTPLCPROPAPIRGFWIVPDILGPTM 480  
 QY 481 ICITSTYECILWPLSTVHPASPPLLTCTREDVEAESRLVLAETPDSFEKHIRSTLQRS 540  
 DB 481 ICITSTYECILWPLSTVHPASPPLLTCTREDVEAESRLVLAETPDSFEKHIRSTLQRS 540  
 QY 541 VANPAFLKASEKDIAPPPCECLQLLSRATQVFREQYILKQDLAKBEIQRRVKLLCDQKKK 600  
 DB 541 VANPAFLKASEKDIAPPPCECLQLLSRATQVFREQYILKQDLAKBEIQRRVKLLCDQKKK 600  
 QY 601 QLEDLSYCREERKSLREMAERLADKYEAKEKQEDIMNMKLLHSHFSELVPLSDSERD 660  
 DB 601 QLEDLSYCREERKSLREMAERLADKYEAKEKQEDIMNMKLLHSHFSELVPLSDSERD 660  
 QY 661 MKKELQLIPDQRLHGLGNAIKQVTMKDYOQKMEKVLSPKPTIILSAYORKCIQSILKE 720  
 DB 661 MKKELQLIPDQRLHGLGNAIKQVTMKDYOQKMEKVLSPKPTIILSAYORKCIQSILKE 720  
 QY 721 EGEHIREMVKQINDIRNHVNF 741

Db 721 EGEHIREMVKQINDIRNHVNF 741

RESULT 2  
 NUP88\_RAT

ID NUP88\_RAT STANDARD; PRT; 742 AA.  
 AC Q08658;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 25-OCT-2004 (Rel. 41, Last sequence update)  
 DE Nuclear pore complex protein Nup88 (Nucleoporin Nup88) (88 kDa nuclear  
 DE pore complex protein) (Nucleoporin Nup84).  
 GN Name=Nup88; Synonyms=Nup84;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND INTERACTION WITH NUP214/CAN.  
 RC TISSUE=Thyroid carcinoma;  
 RX MEDLINE=97311070; PubMed=9166401; DOI=10.1083/jcb.137.5.989;  
 RA Bastos R., Ribas de Pouplana L., Enarson M., Bodoor K., Burke B.;  
 RT "Nup84, a novel nucleoporin that is associated with CAN/Nup214 on the  
 RT cytoplasmic face of the nuclear pore complex.";  
 RL J. Cell Biol. 137:989-1000(1997).  
 CC -!- FUNCTION: Essential component of nuclear pore complex.  
 CC -!- SUBUNIT: Interacts with NUP214/CAN.  
 CC -!- SUBCELLULAR LOCATION: Nuclear pore complex.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U93692; AAB52419.1; -;  
 KW Coiled coil; Nuclear protein; Protein transport; Transport.  
 FT DOMAIN 597 652  
 SQ SEQUENCE 742 AA; 83585 MW; E2C38D0DCBFF3D8 CRC64;

Query Match 89.1%; Score 3461.5; DB 1; Length 742;  
 Best Local Similarity 88.2%; Pred. No. 1.1e-185;  
 Matches 655; Conservative 37; Mismatches 48; Indels 3; Gaps 2;

QY 1 MAAAGPVGDEGELWQTLNPNHVVFLRLREGLNQSPTEAEKPASSSLPSSPPQLLTR 58  
 DB 1 MAAAGPVGDEGELWQTLNPNHVVFLRLREGLNQSPTEAEKPASSSLPSSPPQLPTR 60  
 QY 59 NVVFLGGLFLWDGEDSSFLVRLRGSGGGEPPALSOYORLLCINPPLFEIYQVLLSP 118  
 DB 61 NLVFLGGLFLWDAGSAPLVRLRGSGGSEVPPLSQYQRLCINPPLFEIHQVLLSP 120  
 QY 119 TQHVVALIGIGLWLELPKRWGNSEPEGGKSTVNCSTTPVAERFFTSSTLTUKHAAW 178  
 DB 121 TQHVVALIGIGLWLELPKRWGNSEPEGGKSTVNCSTTPVAERFFTSSTLTUKHAAW 180  
 QY 179 YPSEILDPHVLLTSDNVIIRIYSLREPOTPTNVIILSEAEESLVNKGRTAYTASLGATA 238  
 DB 181 YPSEILDPHVLLTSDNVIIRIYSLREPOTPTNVIILSEAEESLVNKGRTAYTASLGATA 240  
 QY 239 VAPDFGGLDVPKTLFGQNGKDEVVAYPLYLYENGETFLYISLLHSPGNWKAAGVIA 298  
 DB 241 VAPDFGGLDVPKTLFGQNGKDEVVAYPLYLYENGETFLYISLLHSPGNWKAAGVIA 299  
 QY 299 HASAEDNNGYDACAVALCLPCVPIILVIATBSGMLYHCVCVLEGEEDDHTSEKSWDSRID 358  
 DB 300 MHPAEDNNGYDACAVALCLPCVPIILVIATBSGMLYHCVCVLEGEEDDHTSEKSWDSRID 359  
 QY 359 LIPSLYVFECEVELELALKASGEDDPFSDPSCPVKLHRDPKCPSRHYCTHEAGVHSVGL 418

T14156  
kinesin-related protein - African clawed frog  
C/Species: Xenopus laevis (African clawed frog)  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: T14156  
R/Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.  
Cell 91, 357-366, 1997  
A/Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome segregation  
A/Reference number: Z17893; MUID:98028574; PMID:9363944  
A/Accession: T14156  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-2954 <WOO>



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2005, 18:16:00 ; Search time 162 Seconds  
(without alignments)  
1775.601 Million cell updates/sec

Title: US-09-684-890A-2

Perfect score: 3887

Sequence: 1 MAABGPGVGGELWQTLN.....GHIHREVMVQINDIRNHNVP 741

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1736639 seqs, 388188149 residues

Total number of hits satisfying chosen parameters: 1736639

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	285	7.3	62	9 US-09-864-761-40109	Sequence 40109, A
2	201	5.2	40	9 US-09-864-761-39809	Sequence 39809, A
3	194	5.0	858	16 US-10-437-963-140355	Sequence 140355,
4	170	4.4	610	15 US-10-424-598-279883	Sequence 279883,
5	166	4.3	980	15 US-10-211-462-145	Sequence 145, App
6	163	4.2	787	15 US-10-104-047-3340	Sequence 3340, App
7	155	4.0	849	15 US-10-276-774-2414	Sequence 2414, App
8	152	3.9	870	15 US-10-424-599-279884	Sequence 279884,
9	149.5	3.9	849	16 US-10-437-963-194984	Sequence 194984,
10	147.5	3.8	978	17 US-10-732-923-3309	Sequence 3309, App
11	142.5	3.7	1196	15 US-10-282-122A-52737	Sequence 52737, A

12	142.5	3.7	2503	17	US-10-828-985A-11	Sequence 11, Appl
13	142.5	3.7	2543	17	US-10-828-985A-9	Sequence 9, Appl
14	142.5	3.7	2568	17	US-10-828-985A-7	Sequence 7, Appl
15	142.5	3.7	2663	16	US-10-723-860-749	Sequence 749, App
16	140	3.6	1130	15	US-10-260-708-67	Sequence 67, Appl
17	140	3.6	2099	16	US-10-723-860-267	Sequence 267, App
18	140	3.6	2390	15	US-10-092-900A-292	Sequence 292, App
19	139.5	3.6	1293	15	US-10-334-143-2	Sequence 2, Appl
20	138.5	3.6	696	16	US-10-408-765A-1067	Sequence 1067, App
21	138	3.6	1169	15	US-10-369-493-1095	Sequence 1095, App
22	138	3.6	1875	15	US-10-369-493-22285	Sequence 22285, A
23	138	3.6	1875	17	US-10-732-923-3334	Sequence 3334, App
24	138	3.6	1875	17	US-10-732-923-3335	Sequence 3335, App
25	137.5	3.5	798	14	US-10-032-585-7126	Sequence 7126, App
26	137.5	3.5	882	17	US-10-732-923-3300	Sequence 3300, App
27	137.5	3.5	1080	15	US-10-369-493-12560	Sequence 12560, A
28	137	3.5	852	17	US-10-732-923-3313	Sequence 3313, App
29	137	3.5	1955	17	US-10-732-923-3349	Sequence 3349, App
30	136.5	3.5	630	15	US-10-205-647A-6	Sequence 6, Appl
31	136.5	3.5	631	10	US-09-978-309A-48	Sequence 48, Appl
32	136.5	3.5	631	17	US-10-892-831-48	Sequence 48, Appl
33	136.5	3.5	745	15	US-10-108-260A-3547	Sequence 3547, App
34	135.5	3.5	623	16	US-10-416-330-38	Sequence 38, Appl
35	135.5	3.5	798	17	US-10-741-849-7169	Sequence 7169, App
36	135.5	3.5	886	15	US-10-369-493-1016	Sequence 1016, App
37	135.5	3.5	886	17	US-10-732-923-3310	Sequence 3310, App
38	135.5	3.5	1178	15	US-10-282-122A-52434	Sequence 52434, A
39	134	3.4	1979	14	US-10-205-823-419	Sequence 419, App
40	134	3.4	3113	16	US-10-723-860-862	Sequence 862, App
41	134	3.4	3114	16	US-10-751-736-99	Sequence 99, Appl
42	132.5	3.4	1128	17	US-10-732-923-3317	Sequence 3317, App
43	132.5	3.4	2265	15	US-10-092-900A-296	Sequence 296, App
44	132.5	3.4	2327	15	US-10-092-900A-294	Sequence 294, App
45	131.5	3.4	722	16	US-10-437-963-176171	Sequence 176171,

#### ALIGNMENTS

#### RESULT 1

US-09-864-761-40109  
; Sequence 40109 Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40109
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004148.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.4
; OTHER INFORMATION: EST_HUMAN HIT: B5155230.1, EVALUE 1.00e-32
; OTHER INFORMATION: SWISSPROT HIT: O091175, EVALUE 1.50e+00
;
US-09-864-761-40109
```

```

Query Match 7.3%; Score 285; DB 9; Length 62;
Best Local Similarity 85.5%; Pred. No. 4e-15;
Matches 53; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 287 PGNIKAVGSIHAASAAEDNYGDACVLCPCPNILVATESGMLYHCVVLGESEDD 346
||| : : |||
Db 1 PGNIGKLGPLPMHPAAEDNYGDACVLCPCPNILVATESGMLYHCVVLGESEDD 60

QY 347 HT 348
||
Db 61 HT 62
```

```

RESULT 2
US-09-864-761-39809
; Sequence 39809, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39809
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004148.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
; OTHER INFORMATION: EST_HUMAN HIT: AA488609.1, EVALUE 1.00e-15
;
US-09-864-761-39809
```

```

Query Match 5.2%; Score 201; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 550 SEKDIAPPEECQLLSRATQVFRQYILKQDLAKEEIOR 589
Db 1 SEKDIAPPEECQLLSRATQVFRQYILKQDLAKEEIOR 40
```

```

RESULT 3
US-10-437-963-140355
; Sequence 140355, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2005, 18:02:36 ; Search time 44 Seconds  
(without alignments)  
1257.158 Million cell updates/sec

Title: US-09-684-890A-2

Perfect score: 3887

Sequence: 1 MAAAGPVGDELWQTLWLPN.....GEHIREMVKQINDIRHVNPF 741

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3817	98.2	745	4	US-09-949-016-9976
2	156	4.0	192	4	US-09-270-767-44575
3	149.5	3.8	2954	4	US-09-150-867-1
4	142.5	3.7	2662	4	US-09-595-684B-31
5	142.5	3.7	2663	4	US-09-538-092-1252
6	137.5	3.5	1780	4	US-09-949-016-6899
7	137.5	3.5	1786	4	US-09-949-016-7880
8	134	3.4	606	4	US-08-477-831C-2
9	134	3.4	1979	4	US-09-949-016-6468
10	134	3.4	2047	4	US-09-949-016-7404
11	134	3.4	2482	1	US-08-328-254-6
12	134	3.4	3210	4	US-09-538-092-1154
13	134	3.4	3248	1	US-08-353-700-1
14	134	3.4	3248	5	PCT-US95-16216-1
15	131.5	3.4	976	3	US-09-104-324B-4
16	131.5	3.4	976	4	US-09-538-092-1339
17	131	3.4	631	4	US-08-477-831C-11
18	128	3.3	514	2	US-08-960-022-14
19	127.5	3.3	1972	4	US-08-875-435B-3
20	126.5	3.3	1886	3	US-08-938-105-3
21	126	3.2	1055	4	US-09-949-016-9776
22	125.5	3.2	2285	3	US-09-308-375-2
23	125.5	3.2	2285	4	US-09-932-183A-2
24	125	3.2	630	4	US-09-248-796A-20275
25	125	3.2	1201	3	US-09-098-901-2
26	123	3.2	994	4	US-09-949-016-6779
27	122.5	3.2	592	2	US-08-736-770-6

28 122.5 3.2 592 4 US-09-702-705-1809 Sequence 1809, Ap  
29 122.5 3.2 592 4 US-09-736-457-1809 Sequence 1809, Ap  
30 122.5 3.2 592 4 US-09-643-657-4 Sequence 4, Appli  
31 122.5 3.2 592 4 US-09-671-325-1809 Sequence 1809, Ap  
32 122 3.1 1898 1 US-08-056-200-94 Sequence 94, Appl  
33 122 3.1 1898 2 US-08-800-644-94 Sequence 94, Appl  
34 122 3.1 1898 4 US-09-538-092-1280 Sequence 1280, Ap  
35 122 3.1 1972 4 US-08-875-435B-4 Sequence 4, Appli  
36 121.5 3.1 605 4 US-09-949-016-8823 Sequence 8823, Ap  
37 121.5 3.1 1799 4 US-09-134-000C-5178 Sequence 5178, Ap  
38 121 3.1 589 4 US-09-643-657-14 Sequence 14, Appl  
39 121 3.1 631 4 US-09-949-016-11595 Sequence 11595, A  
40 121 3.1 1786 3 US-08-973-462-8 Sequence 8, Appli  
41 120.5 3.1 1088 3 US-09-085-199B-11 Sequence 11, Appl  
42 120 3.1 2568 4 US-09-866-108A-3 Sequence 3, Appli  
43 119.5 3.1 1307 4 US-09-949-016-7561 Sequence 7561, Ap  
44 119.5 3.1 1401 4 US-09-750-590A-2 Sequence 2, Appli  
45 119.5 3.1 1857 4 US-09-917-254-91 Sequence 91, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-949-016-9976

; Sequence 9976, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9976

; LENGTH: 745

; TYPE: PRT

; ORGANISM: Human

; US-09-949-016-9976

Query Match 98.2%; Score 3817; DB 4; Length 745;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 729; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MAAAGPVGDELWQTLWLPNHHVFLRLREGLNQSPTEAEKPASSSLPSSPPOLLTRNV 60  
Db 5 MAAAGPVGDELWQTLWLPNHHVFLRLREGLNQSPTEAEKPASSSLPSSPPOLLTRNV 64  
Qy 61 VFGLGELFLWDGEDSSFLVRLRPGSGGEEPALSOYQLLCLINPFLFIYQVLLSPTQ 120  
Db 65 VFGLGELFLWDGEDSSFLVRLRPGSGGEEPALSOYQLLCLINPFLFIYQVLLSPTQ 124  
Qy 121 HHVALIGIKLWLELPKRWGNSEFEGGKSTVNCSTTPVAERFPTSSTSLTLKHAAYP 180  
Db 125 HHVALIGIKLWLELPKRWGNSEFEGGKSTVNCSTTPVAERFPTSSTSLTLKHAAYP 184  
Qy 181 SETLDHVVLLTSDNVIIRIYSLREPOTPTNVIILSEAEESVLNKGRAVTASIGETAVA 240  
Db 185 SETLDHVVLLTSDNVIIRIYSLREPOTPTNVIILSEAEESVLNKGRAVTASIGETAVA 244  
Qy 241 FDFGGLDVAPEKTLFGQNGKDEVVAYPLIYENGETFLTYSLLHSPGNWKVAVGSIHAH 300  
Db 245 FDFGGLAAPVPEKTLFGQNGKDEVVAYPLIYENGETFLTYSLLHSPGNKGLGLPLPMH 304  
Qy 301 SAAEDNYGDACAVLCPLPCVFNILVIATBSGMLYHCVVLEGEEDDDHSTSEKSDSRIDL 360

```

Db 305 PAAEDNYGDACAVLCLPVPNIVLATSGMLYHCVLEGEDEDDHTSEKSWDSRIDLI 364
QY 361 PSILYFECVELELALKLAGEDDPDSFSCPKLHRDPKPSRVHCTHEAGVHSGLTW 420
Db 365 PSILYFECVELELALKLAGEDDPDSFSCPKLHRDPKPSRVHCTHEAGVHSGLTW 424
QY 421 IHKLHKLFGSDEEDKDSLOELSTEQKCFVEHILCTRPFCROPAPIRGFWIVDPDILGPTM 480
Db 425 IHKLHKLFGSDEEDKDSLOELSTEQKCFVEHILCTRPFCROPAPIRGFWIVDPDILGPTM 484
QY 481 ICITSTYECILWPLISTVHPASPPLCTREDVEVAESSLRVLAETPDSFEKHRSILQRS 540
Db 485 ICITSTYECILWPLISTVHPASPPLCTREDVEVAESSLRVLAETPDSFEKHRSILQRS 544
QY 541 VANPAFLKASEKDIAAPPPECQLLSRATQVFEQVILKODLAKBEIQRVXKLCDOKKK 600
Db 545 VANPAFLKASEKDIAAPPPECQLLSRATQVFEQVILKODLAKBEIQRVXKLCDOKKK 604
QY 601 QLEDLSYCREERKSREMAERLADKYEEAKEQEDIMNRMKLLHSFHSFSELVLSDSERD 660
Db 605 QLEDLSYCREERKSREMAERLADKYEEAKEQEDIMNRMKLLHSFHSFSELVLSDSERD 664
QY 661 MKKELQILPDQLRHGLGNAIKQVTMKDYQQQMKVKVLSLPKPTIILSAYORKCISILKE 720
Db 665 MKKELQILPDQLRHGLGNAIKQVTMKDYQQQMKVKVLSLPKPTIILSAYORKCISILKE 724
QY 721 EGEHIREMYKQINDIRNHVNF 741
Db 725 EGEHIREMYKQINDIRNHVNF 745

```

RESULT 2

```

US-09-270-767-44575
; Sequence 44575, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44575
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44575

```

```

Query Match 4.0%; Score 156; DB 4; Length 192;
Best Local Similarity 27.4%; Pred. No. 3.5e-06;
Matches 46; Conservative 40; Mismatches 74; Indels 8; Gaps 4;

QY 574 EGYLLKQDLAKBEIORRVKLLCDQKKQLEDLSYCREERKSREMAERLADKYEEAKEQ 633
Db 1 EGYLLKQDLAKBEIORRVKLLCDQKKQLEDLSYCREERKSREMAERLADKYEEAKEQ 60
QY 634 EDIMNRMKLLHSFHSFSELVLSDSERDMMKELQLIPDQLRHGLGNAIKQVTMKDYQQQRM 693
Db 61 ELLVRKCNALMQRANASLPNSVIAEREFQSVIRLNKVTQSLAAGLE--TAKTFTNKQRY 118
QY 694 EKVLG---LPKPTIILSAYORKCISILKE--EGEHIEM--VKQINDI 735
Db 119 HIAQSQEDLKNAYELPEKQHRITTEILTLQAGEIDRQITDKRINKI 166

```

RESULT 3

```

US-09-150-867-1
; Sequence 1, Application US/09150867
; Patent No. 6645748

```

```

; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth W.
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S.B.
; APPLICANT: Cleveland, Don W.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for
; FILE REFERENCE: 18557C-000110US
; CURRENT APPLICATION NUMBER: US/09/150,867
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: US 60/058,645
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2954
; TYPE: PRT
; ORGANISM: Xenopus sp.
; FEATURE:
; OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)
; OTHER INFORMATION: member of the kinesin superfamily of microtubule
; OTHER INFORMATION: motor proteins
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(472)
; OTHER INFORMATION: kinesin like motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (473)..(2752)
; OTHER INFORMATION: rod domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (2753)..(2954)
; OTHER INFORMATION: tail domain
; OTHER INFORMATION: tail domain
US-09-150-867-1

Query Match 3.8%; Score 149.5; DB 4; Length 2954;
Best Local Similarity 18.5%; Pred. No. 0.0013;
Matches 74; Conservative 81; Mismatches 157; Indels 89; Gaps 14;

QY 370 EUELAKLAIASGDDPDSFSCPKLHRDPKPSRVHCTHEAGVHSGLTWIKLHKFLG 429
Db 1581 ELQALAKLAIASGDDPDSFSCPKLHRDPKPSRVHCTHEAGVHSGLTWIKLHKFLG 429
QY 430 SDEE---DKDSLOELSTEQKCFVEHILCTRPFCROPAPIRGFWIVDPDILGPTMCIITST 486
Db 1629 EQEKLHKEKLEQAEQVLEKCEVHLM-----KSMIESKSS 1664
QY 487 YECLIWPLISTVHPASPPLCTREDVEVAESSLRVLAETPDSFEKHRSI-----LQRSV 541
Db 1665 LE---SLQHEKHDTEQQLLAKQMQVVTQEKELQOQTHEULTAEVDHLKENIELGLNF 1720
QY 542 ANPAFLKASEKDIAAPPPECQLLSRATQVFEQVILKODLAKBEIQRVXKLCDOKKKQ 601
Db 1721 KNEAQKTKT-----EQC--LLNENKELEQSHRLOCEI--EELMKSLK----DKESA 1765
QY 602 LEDLSYCREERKSREMAERLADKYEEAKEQEDIMNRMKLLHSFHSFSELVLSDSERDM 661
Db 1766 LETLKESEQKVINLQNMVLEMBELKNSORTVIAERDQLQDLRESVMSIETQDDL 1825
QY 662 KK-----ELQLIPDQLRHGLGNAIKQVTMKDYQQQKMEKYL 697
Db 1826 RAAQALQOKQKQVLETSQISVLEKISLLENQMLYNAVTVKETSERDDLQNSQKHFLF 1885
QY 698 SLPKPTIILSAYORK-CIOSILKEGEHIREMYKQINDIRN 737
Db 1886 S-EIETLSLSLKEFALEQAEKADAAARTIDITEKISN 1925

```

RESULT 4

```

US-09-595-684B-31
; Sequence 31, Application US/09595684B

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2005, 17:47:00 ; Search time 169 Seconds  
(without alignments)  
1695.796 Million cell updates/sec

Title: US-09-684-890A-2

Perfect score: 3887

Sequence: 1 MAAGGPGVGDGLMQLP.....GSHIREMVQINDIRHNVNF 741

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3887	100.0	741	AAE01161	Aae01161 Human Nup
2	3874	99.7	741	AAAM39226	Aam39226 Human pol
3	3871	99.6	757	AAAM79827	Aam79827 Human pro
4	3871	99.6	757	AAAM79826	Aam79826 Human pro
5	3871	99.6	757	AAAM41012	Aam41012 Human pol
6	3817	98.2	741	ADK60214	Adk60214 Angiogene
7	3817	98.2	741	ADK60515	Adk60515 Angiogene
8	3817	98.2	741	ADP71138	Adp71138 Angiogene
9	3817	98.2	741	ABM81818	Abm81818 Tumour-as
10	3461.5	89.1	742	AAE01162	Aae01162 Rat Nup88
11	2313	59.5	443	AAAM78842	Aam78842 Human pro
12	1561	40.2	303	AAAM78843	Aam78843 Human pro
13	648.5	16.7	702	AAE01160	Aae01160 Drosophil
14	646.5	16.6	702	ABBT71509	Abbt71509 Drosophil
15	285	7.3	62	AAAM20196	Aam20196 Peptide #
16	285	7.3	62	ABBA40469	Abba40469 Peptide #
17	285	7.3	62	AAAM34190	Aam34190 Peptide #
18	285	7.3	62	ABB24811	Abb24811 Protein #
19	285	7.3	62	AAAM74021	Aam74021 Human bon
20	285	7.3	62	AAAM61293	Aam61293 Human bon
21	285	7.3	62	ABGS5775	Abgs5775 Human liv
22	285	7.3	62	ABGS43915	Abgs43915 Human liv
23	245.5	6.3	822	AAAG41929	Aag41929 Arabidops
24	219.5	5.6	767	AAAG41930	Aag41930 Arabidops
25	211.5	5.4	725	AAAG41931	Aag41931 Arabidops

26	201	5.2	40	4	AAAM19951	Aam19951 Peptide #
27	201	5.2	40	4	ABB39985	Abb39985 Peptide #
28	201	5.2	40	4	AAAM33606	Aam33606 Peptide #
29	201	5.2	40	4	ABB24511	Abb24511 Protein #
30	201	5.2	40	4	AAAM73403	Aam73403 Human bon
31	201	5.2	40	4	AAAM60732	Aam60732 Human bra
32	201	5.2	40	4	ABGS5125	Abgs5125 Human liv
33	201	5.2	40	5	ABG43262	Abg43262 Human pep
34	166	4.3	980	6	ABU03527	Abu03527 Angiogene
35	166	4.3	980	8	ADN03810	Adn03810 Antipsori
36	166	4.3	980	8	ADP23344	Adp23344 PRO poly
37	163	4.2	787	7	ADSG5186	Adsg5186 Human pro
38	163	4.2	989	8	ADS88434	Ads88434 Human pro
39	158	4.1	296	4	AAAB94454	Aaab94454 Human pro
40	155	4.0	344	4	AAAM79051	Aam79051 Human pro
41	155	4.0	370	4	ABB12044	Abb12044 Human ret
42	155	4.0	370	4	AAAM80035	Aam80035 Human pro
43	155	4.0	609	7	ADM26165	Adm26165 Hyperther
44	149.5	3.8	2954	2	AAAY01632	Aay01632 Amino aci
45	143	3.7	766	4	ABU53068	Abu53068 Transport

#### ALIGNMENTS

##### RESULT 1

AAE01161

ID AAE01161 standard; protein; 741 AA.

XX AAE01161;

DT 17-JUL-2001 (first entry)

DE Human Nup88 homology protein.

XX Transcription factor-selective nuclear transport receptor; dorsal;

KW nuclear factor-kappa B; NF-kB; nuclear pore; humoral response; therapy;

KW inflammation; asthma; rheumatoid arthritis; RA; septic shock;

KW chronic obstructive pulmonary disease; COPD; angiogenesis; lung fibrosis;

KW glomerulonephritis; atherosclerosis; apoptotic disorder; antibacterial;

KW acquired immune deficiency syndrome; AIDS; immunosuppressive;

KW nephrotropic; human; Nup88.

XX Homo sapiens.

OS Homo sapiens.

XX WO200129087-A1.

XX 26-APR-2001.

PD 19-OCT-2000; 2000WO-SE002022.

XX 22-OCT-1999; 99SE-00003832.

XX (INNA-) INNATE PHARM AB.

PA Samakovlis C, Uv AE;

XX WPI; 2001-290899/30.

XX Novel purified transcription factor-selective nuclear transport receptor

PT polypeptide is used to treat, prevent and diagnose inflammation, asthma,

PT rheumatoid arthritis, atherosclerosis, AIDS glomerulonephritis and

PT apoptotic disorders.

XX Claim 6; Page 35-37; 42pp; English.

PS The present sequence is human Nup88 protein which localise to the  
CC cytoplasmic filaments of the nuclear pores to provide binding sites for  
CC nuclear import substrates. The Nup88 protein has homology to Drosophila  
CC transcription factor-selective nuclear transport receptor. The  
CC transcription factor-selective nuclear transport receptor mediates the  
CC import of dorsal nuclear factor-kappa B (NF-kB) protein at the level of  
CC the nuclear pore and is required for activation of the Drosophila humoral

CC immune response. It is also used for prevention, diagnosis and treatment  
CC of pathophysiological disorders related to the family of nuclear  
CC receptors such as inflammation, asthma, rheumatoid arthritis (RA),  
CC chronic obstructive pulmonary disease (COPD), angiogenesis, septic shock,  
CC lung fibrosis, glomerulonephritis, atherosclerosis, AIDS and apoptotic  
CC disorders. It is also used in screening assays to identify its agonists  
CC or inhibitors  
XX  
SQ Sequence 741 AA;

Query Match 100.0%; Score 3887; DB 4; Length 741;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAGPVGDELWQTLNHNHVFRLRGLKQNSPTEAEKPASSSLPSPQLLTRNV 60  
DB 1 MAAAGPVGDELWQTLNHNHVFRLRGLKQNSPTEAEKPASSSLPSPQLLTRNV 60  
QY 61 VFGLGELFLWDGEDSSFLVRLRGPSSGGEPPALSYQVRLCINPPLFEIYQVLLSPTQ 120  
DB 61 VFGLGELFLWDGEDSSFLVRLRGPSSGGEPPALSYQVRLCINPPLFEIYQVLLSPTQ 120  
QY 121 HHVALIGIKGLMWLELPKWKGNSEPEGGKSTVNCSTTPVAERFTSSTSLTLKHAAYTP 180  
DB 121 HHVALIGIKGLMWLELPKWKGNSEPEGGKSTVNCSTTPVAERFTSSTSLTLKHAAYTP 180  
QY 181 SETLDPHVLLTSDNVIRIYSREPTQTNVILSEAEESVLNKGRAVYASLGATA 240  
DB 181 SETLDPHVLLTSDNVIRIYSREPTQTNVILSEAEESVLNKGRAVYASLGATA 240  
QY 241 FDFGGLDAPVKTLFGQNGKDEWVAYPLYLYENGETFLTYISLLHSPGNKXAVGSIHA 300  
DB 241 FDFGGLDAPVKTLFGQNGKDEWVAYPLYLYENGETFLTYISLLHSPGNKXAVGSIHA 300  
QY 301 SAAEDNYGDACAVLCLCPVNLIVTATSGMLYHCVLGESEDDHTSEKSWDSRIDLI 360  
DB 301 SAAEDNYGDACAVLCLCPVNLIVTATSGMLYHCVLGESEDDHTSEKSWDSRIDLI 360  
QY 361 PSLYVFECELELALKLAGEDDPDSFSCPKLHRDPCPSRVHCTHEAGVHSLGTW 420  
DB 361 PSLYVFECELELALKLAGEDDPDSFSCPKLHRDPCPSRVHCTHEAGVHSLGTW 420  
QY 421 IHKLHKLGSDEEDKDSLOELSTEQCFVEHILCTPLPCROPAPIRGFWIVPDILGPTM 480  
DB 421 IHKLHKLGSDEEDKDSLOELSTEQCFVEHILCTPLPCROPAPIRGFWIVPDILGPTM 480  
QY 481 ICITSTYECLIWPLISTVHPASPLLCREDVEVAESSLRVLAETPDSPEKHRSILQRS 540  
DB 481 ICITSTYECLIWPLISTVHPASPLLCREDVEVAESSLRVLAETPDSPEKHRSILQRS 540  
QY 541 VANPAFLKASEKDIAAPPPECQLLSRATQVFEQVILKODIAKEIQRVVKLLCDOKKK 600  
DB 541 VANPAFLKASEKDIAAPPPECQLLSRATQVFEQVILKODIAKEIQRVVKLLCDOKKK 600  
QY 601 QLEDLSYCREERSKUREMAERLADKYEEAKEQEDIMNMKKLLHSFHSFSLVLSDSERD 660  
DB 601 QLEDLSYCREERSKUREMAERLADKYEEAKEQEDIMNMKKLLHSFHSFSLVLSDSERD 660  
QY 661 MKKELQILPDQLRHGNAIKQVTMKDYQOQKMEKVLSPKPTIILSAYQRCIQSILKE 720  
DB 661 MKKELQILPDQLRHGNAIKQVTMKDYQOQKMEKVLSPKPTIILSAYQRCIQSILKE 720  
QY 721 EGEHIREMVKQINDIRNHVNF 741  
DB 721 EGEHIREMVKQINDIRNHVNF 741

RESULT 2  
ID AAM39226  
AC AAM39226 standard; protein; 741 AA.  
XX AAM39226;  
XX

DT 22-OCT-2001 (first entry)  
XX Human polypeptide SEQ ID NO 2371.  
XX  
KW Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
PR 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
Zhou P, Goodrich R, Drmanac RT;  
WPI; 2001-442253/47.  
DR N-PSDB; AAI58382.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX  
PS Example 4; SEQ ID NO 2371; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with nototropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 741 AA;

Query Match 99.7%; Score 3874; DB 4; Length 741;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 739; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAAGPVGDELWQTLNHNHVFRLRGLKQNSPTEAEKPASSSLPSPQLLTRNV 60  
DB 1 MAAAGPVGDELWQTLNHNHVFRLRGLKQNSPTEAEKPASSSLPSPQLLTRNV 60  
QY 61 VFGLGELFLWDGEDSSFLVRLRGPSSGGEPPALSYQVRLCINPPLFEIYQVLLSPTQ 120  
DB 61 VFGLGELFLWDGEDSSFLVRLRGPSSGGEPPALSYQVRLCINPPLFEIYQVLLSPTQ 120

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 06:10:34 ; Search time 7720 Seconds  
(without alignments)  
11724.978 Million cell updates/sec

Title: US-09-684-890A-1  
Perfect score: 2378  
Sequence: 1 gataaacccacagacacaa.....tataaaagggttttgatg 2378

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1685	70.9	2430	3 AK02643	Mus muscu
2	1671.2	70.3	2452	3 AK028563	Mus muscu
3	1552.6	65.3	2076	3 AK085776	Mus muscu
4	1220.8	51.3	3004	3 AK050893	Mus muscu
5	1014.8	42.7	1978	3 AK039007	Mus muscu
c	6	918.8	38.6	954 1	AL525959 Mus muscu
7	894.6	37.6	935	1 AL526005	AL526005 AL526005
8	865.6	36.4	947	5 BU193560	BU193560 AGENCOURT
9	859.6	36.1	1096	5 BM923340	BM923340 AGENCOURT
10	855	36.0	1149	4 BM541613	BM541613 AGENCOURT
11	837	35.2	921	5 BU526916	BU526916 AGENCOURT
c	12	823	34.6	932 5	BX349939 BX349939
13	820.2	34.5	897	5 BX349940	BX349940 BX349940
c	14	780.8	32.8	945 5	BQ953718 AGENCOURT
15	778	32.7	887	6 CD359338	CD359338 AGENCOURT
16	769.2	32.3	843	4 BG759661	BG759661 602713354
17	764	32.1	866	5 BU858009	BU858009 AGENCOURT
18	762.2	32.1	902	5 BU539430	BU539430 AGENCOURT
19	761.2	32.0	789	4 BI819253	BI819253 603034680
20	758	31.9	921	5 BQ647182	BQ647182 AGENCOURT
21	755.2	31.8	833	4 BI552947	BI552947 603197875
22	742.4	31.2	873	4 BG219184	BG219184 RST38937
23	737.6	31.0	762	4 BG686994	BG686994 602650990
24	735.2	30.9	795	6 CB997327	CB997327 AGENCOURT

#### RESULT 1

AK002643

LOCUS

DEFINITION

AK002643

AK002643

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

#### ALIGNMENTS

RESULT 1  
LOCUS AK002643 2430 bp mRNA linear HTC 03-APR-2004  
DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610016P04 product:preimplantation protein 2, full insert sequence.  
ACCESSION AK002643  
VERSION AK002643.1 GI:112832779  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE 5  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

25 734.8 30.9 816 4 BI549458  
26 733.4 30.8 832 4 BI461310  
27 732.6 30.8 831 5 BU558277  
c 28 727 30.6 807 4 BI194807  
29 719.4 30.3 799 2 BE902870  
30 717.6 30.2 740 7 CF122933  
31 715.2 30.1 796 4 BG187429  
c 32 709 29.8 752 5 BQ182096  
c 33 704.8 29.6 756 1 AA868783  
34 700.2 29.4 909 6 CD557705  
35 698 29.4 816 4 BI458058  
36 697.8 29.3 784 4 BG686048  
37 697.8 29.3 792 6 CB990577  
38 697.4 29.3 839 4 BI598624  
39 693 29.1 718 4 BG720209  
40 692.6 29.1 798 4 BI550338  
41 690.6 29.0 832 5 BU940824  
42 682.8 28.7 783 4 BI459146  
43 680.4 28.6 690 4 BG719270  
44 677.4 28.5 730 4 BG718589  
45 673.8 28.3 965 4 BI461277

BI549458 603192123  
BI461310 603207250  
BU558277 AGENCOURT  
BI194807 RST14108  
BE902870 601675861  
CF122933 UI-HF-CB0  
BG187429 RST6420 A  
BQ182096 UI-H-EU0-  
AA868783 al49e02.s  
CD557705 AGENCOURT  
BI458058 603190069  
BG686048 602638617  
CB990577 AGENCOURT  
BI598624 603251377  
BG720209 602692257  
BI550338 603196275  
BU940824 AGENCOURT  
BI459146 603193125  
BG719270 602690339  
BG718589 602696675  
BI461277 603207215



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 11:18:28 ; Search time 1482 Seconds  
(without alignments)  
10179.925 Million cell updates/sec

Title: US-09-684-890A-1  
Perfect score: 2378  
Sequence: 1 gataaacacacagacacaa.....tataaaagggtgtttgatg 2378

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 717243 seqs, 3172129809 residues

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2341.6	98.5	2392	15	US-10-037-270-261 Sequence 261, App
2	2341.6	98.5	2392	17	US-10-117-722-261 Sequence 261, App
3	1710.2	71.9	2229	18	US-10-261-175A-18 Sequence 18, Appl
c	1411.6	59.4	8263	17	US-10-062-674-1762 Sequence 1762, Ap
5	393.4	16.5	456	10	US-09-918-995-15834 Sequence 15834, A
6	332	14.0	356	9	US-09-796-692-2684 Sequence 2684, Ap
7	332	14.0	356	14	US-10-040-862-2684 Sequence 2684, Ap

8	332	14.0	356	17	US-10-057-475B-2684	Sequence 2684, Ap
9	332	14.0	356	17	US-10-154-884B-2684	Sequence 2684, Ap
10	332	14.0	356	19	US-10-764-324-2684	Sequence 2684, Ap
11	322.2	13.5	429	10	US-09-918-995-8447	Sequence 8447, Ap
12	243.4	10.2	514	16	US-10-029-386-2600	Sequence 2600, Ap
13	182.8	7.7	493	9	US-09-783-590-9025	Sequence 9025, Ap
C 14	166	7.0	466	9	US-09-864-761-6544	Sequence 6544, Ap
C 15	165.4	7.0	187	9	US-09-864-761-23256	Sequence 23256, Ap
C 16	146.4	6.2	212	9	US-09-933-797-642	Sequence 642, App
C 17	145	6.1	147	16	US-10-029-386-16300	Sequence 16300, A
C 18	134.6	5.7	459	9	US-09-864-761-6262	Sequence 6262, Ap
C 19	134.6	5.7	733	13	US-10-027-632-127047	Sequence 127047, Ap
C 20	134.6	5.7	733	17	US-10-027-632-127047	Sequence 127047, Ap
C 21	134.6	5.7	831	13	US-10-027-632-127046	Sequence 127046, Ap
C 22	134.6	5.7	831	17	US-10-027-632-127046	Sequence 127046, Ap
C 23	133.8	5.6	420	17	US-10-062-674-816	Sequence 816, App
C 24	122	5.1	122	9	US-09-864-761-22981	Sequence 22981, A
C 25	105.6	4.4	543	16	US-10-029-386-13044	Sequence 13044, A
C 26	103.4	4.3	112	16	US-10-029-386-26744	Sequence 26744, A
C 27	84	3.5	805	13	US-10-027-632-127043	Sequence 127043, A
C 28	84	3.5	805	13	US-10-027-632-127044	Sequence 127044, A
C 29	84	3.5	805	13	US-10-027-632-127045	Sequence 127045, A
C 30	84	3.5	805	13	US-10-027-632-154416	Sequence 154416, A
C 31	84	3.5	805	17	US-10-027-632-127043	Sequence 127043, A
C 32	84	3.5	805	17	US-10-027-632-127044	Sequence 127044, A
C 33	84	3.5	805	17	US-10-027-632-127045	Sequence 127045, A
C 34	84	3.5	805	17	US-10-027-632-154416	Sequence 154416, A
C 35	60	2.5	60	10	US-09-908-975-10796	Sequence 10796, A
C 36	54.4	2.3	65	10	US-09-908-975-27778	Sequence 27778, A
C 37	46	1.9	715	20	US-10-425-115-24320	Sequence 24320, A
C 38	45	1.9	1702	18	US-10-302-172-574	Sequence 574, App
C 39	44	1.9	694	14	US-10-184-644-60	Sequence 60, Appl
C 40	44	1.9	694	14	US-10-184-634-60	Sequence 60, Appl
C 41	42.6	1.8	414	19	US-10-437-963-56419	Sequence 56419, A
C 42	41.8	1.8	1840	17	US-10-264-049-284	Sequence 284, App
C 43	41.8	1.8	2383	17	US-10-104-047-1967	Sequence 1967, App
C 44	41.8	1.8	2383	21	US-10-887-553A-1206	Sequence 1206, Ap
C 45	41.2	1.7	145068	19	US-10-322-281-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1  
US-10-037-270-261  
; Sequence 261, Application US/10037270  
; Publication No. US20030104529A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyang  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yundong  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Tillinghast, John  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/10/037,270  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21									
; NUMBER OF SEQ ID NOS: 1104									
; SOFTWARE: pt_FL_genes Version 1.0									
; SEQ ID NO 261									
; LENGTH: 2392									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (49)..(2274)									
US-10-037-270-261									
Query Match 98.5%; Score 2341.6; DB 15; Length 2392;									
Best Local Similarity 99.2%; Pred. No. 0;									
Matches 2353; Conservative 0; Mismatches 19; Indels 0; Gaps 0;									
QY	7	CCACAGACACAAACATACCTTTGAGCAGCTTGGGCCAAGATGGCGGCCGCGAGGGA	66	Db	847	CTGTACATCTTATATGAAAAATGGAGAGACTTCTCTGACATACATCAGTCTGTGTACACAGC	906		
Db	7	CCCTCGAGGCCAAGAAATTCGGCACGAGAGCAGTTGGGCCAAGATGGCGGCCGCGAGGGA	66	QY	907	CCTGGAATATTTGGAAGCTGTTGGGTCCATTGGCCCATCTGCGGCTGAAGATAAC	966		
QY	67	CGGGTGGCGGCGGAGCTGTGGCAGACCTGGCTTCTTAAACAGCTGCTGTTCTTCGGG	126	Db	907	CCTGGAATATTTGGAAGCTGTTGGGTCCATTGGCCCATCTGCGGCTGAAGATAAC	966		
Db	67	CGGGTGGCGGCGGAGCTGTGGCAGACCTGGCTTCTTAAACAGCTGCTGTTCTTCGGG	126	QY	967	TATGGTTATGATGCGTGTGCTACTCTGCTTACCTGTGTGCCCAATATCTTAGTGATC	1026		
QY	127	CTCCGGGAGGAGTCAAAAAACAGAGTCCAAACCGAAGCTGAGAAACAGCTTCTTCGTG	186	Db	967	TATGGTTATGATGCGTGTGCTACTCTGCTTACCTGTGTGCCCAATATCTTAGTGATC	1026		
Db	127	CTCCGGGAGGAGTCAAAAAACAGAGTCCAAACCGAAGCTGAGAAACAGCTTCTTCGTG	186	QY	1027	GCTACTGAATCAGGAATGCTGTATCACTGTGTGCTGTAGAAAGGAGGAGAGATGAC	1086		
QY	187	TTGGCTTCTGTCGCGCGCGGCTGTGTGACGAGAAACAGTGGTCTTTGGCCTCGGCGGA	246	Db	1027	GCTACTGAATCAGGAATGCTGTATCACTGTGTGCTGTAGAAAGGAGGAGAGATGAC	1086		
Db	187	TTGGCTTCTGTCGCGCGCGGCTGTGTGACGAGAAACAGTGGTCTTTGGCCTCGGCGGA	246	QY	1087	CACACGTGAGAAAAGTCTCGGGATTCAGGATTTGACCTCATTCCTTCTCTGTATGTGTTT	1146		
QY	247	GAGCTTTTCTGTGGACGGAGACAGACCTCTTCTTGTAGTGTGCTTGGCTTGGGGCCCC	306	Db	1087	CACACGTGAGAAAAGTCTCGGGATTCAGGATTTGACCTCATTCCTTCTCTGTATGTGTTT	1146		
Db	247	GAGCTTTTCTGTGGACGGAGACAGACCTCTTCTTGTAGTGTGCTTGGCTTGGGGCCCC	306	QY	1147	GAATGTTGAGTTCGGAGCTTGTGAACTTGAACCTGGCATCTGGAGAGATGACCTTTTGAT	1206		
QY	307	ACGGCGGGCGGGAAGAGCCCGCTGTCCAGTACACAGAGATTTGCTTTGCATAAATCCA	366	Db	1147	GAATGTTGAGTTCGGAGCTTGTGAACTTGAACCTGGCATCTGGAGAGATGACCTTTTGAT	1206		
Db	307	ACGGCGGGCGGGAAGAGCCCGCTGTCCAGTACACAGAGATTTGCTTTGCATAAATCCA	366	QY	1207	TCTGACTTTTCTTGTCCAGTCAAACCTTCATAGAGATCCCAAGTGTCTTCAAGATATCAC	1266		
QY	367	CCGCTGTTTGAATCTATCAAGTCTTGTGTTAAAGCCCAACACATCATGTAGCACTTATA	426	Db	1207	TCTGACTTTTCTTGTCCAGTCAAACCTTCATAGAGATCCCAAGTGTCTTCAAGATATCAC	1266		
Db	367	CCGCTGTTTGAATCTATCAAGTCTTGTGTTAAAGCCCAACACATCATGTAGCACTTATA	426	QY	1267	TGTACTCATGAAGCTGTGTGTACATAGTGTGGGCTAACTTGGATTCATATAAATCTTACAAA	1326		
QY	427	GGAATTAAGAGGACTTATGTTATAGAAATTTACCTTAAAGATGGGGAAGAAATTTCTGAA	486	Db	1267	TGTACTCATGAAGCTGTGTGTACATAGTGTGGGCTAACTTGGATTCATATAAATCTTACAAA	1326		
Db	427	GGAATTAAGAGGACTTATGTTATAGAAATTTACCTTAAAGATGGGGAAGAAATTTCTGAA	486	QY	1327	TTTCTTGGATCAGATGAAAGAGATAAGGATAGTTTACAGGAACTCTCTACAGAAACAGAAA	1386		
QY	487	GAAAGTGGAAAAATCAACAGTGAATTTGTAGTACCACTCCAGTTGCGGAGAGATTTTTCACC	546	Db	1327	TTTCTTGGATCAGATGAAAGAGATAAGGATAGTTTACAGGAACTCTCTACAGAAACAGAAA	1386		
Db	487	GAAAGTGGAAAAATCAACAGTGAATTTGTAGTACCACTCCAGTTGCGGAGAGATTTTTCACC	546	QY	1387	TGCTTTTGTGAACACATCCTTTGTACGAGGCCATTGCCCCTGCAGGAGCCAGCTCCAATT	1446		
QY	547	AGTTCCACTCTCTGACTTAAAGCATGCTGATGATGATGATGATGATGATGATGATGATGAT	606	Db	1387	TGCTTTTGTGAACACATCCTTTGTACGAGGCCATTGCCCCTGCAGGAGCCAGCTCCAATT	1446		
Db	547	AGTTCCACTCTCTGACTTAAAGCATGCTGATGATGATGATGATGATGATGATGATGATGAT	606	QY	1447	CGAGATTTTGGATTTGATCTGACATTTCTGGGACCCAGAGATCTGCATCACCAGTACC	1506		
QY	607	CACGTAGTCTTAAACATCAGAACCGTAATCAGAAATTTACTCTCTAGTGGCGGAG	666	Db	1447	CGAGATTTTGGATTTGATCTGACATTTCTGGGACCCAGAGATCTGCATCACCAGTACC	1506		
Db	607	CACGTAGTCTTAAACATCAGAACCGTAATCAGAAATTTACTCTCTAGTGGCGGAG	666	QY	1507	TATGAATGCTCATATGCGCGTTATTAAGTACAGTCCATCCAGCGTCTCCTCCCTGCTT	1566		
QY	667	ACACCCACTTAAAGTATATCTTTTTCAGAACCGGAGGAAAGTCTAGTACTCAATAAA	726	Db	1507	TATGAATGCTCATATGCGCGTTATTAAGTACAGTCCATCCAGCGTCTCCTCCCTGCTT	1566		
Db	667	ACACCCACTTAAAGTATATCTTTTTCAGAACCGGAGGAAAGTCTAGTACTCAATAAA	726	QY	1567	TGTACTCGAGAAGATGTTGAAAGTGGCAGAGTCTTCCCTCCGCTGTTCTGGCTGAAACCCCA	1626		
QY	727	GGAAGGGGCTATACCGCATCTCTAGAGAGACAGCAGTTGCAATTTGATTTTGGGCCATTG	786	Db	1567	TGTACTCGAGAAGATGTTGAAAGTGGCAGAGTCTCCCTCCGCTGTTCTGGCTGAAACCCCA	1626		
Db	727	GGAAGGGGCTATACCGCATCTCTAGAGAGACAGCAGTTGCAATTTGATTTTGGGCCATTG	786	QY	1627	GATTCCTTTGAAAAGCATATTTAGAGAGCATTTTGAACGTAAGTGTGTTGCCAATCCAGCATTT	1686		
QY	787	GACGAGTCCCAAGACTCTATTTTGGACAAAACGCGAAGATGAAGTGTGCGCATACCCA	846	Db	1627	GATTCCTTTGAAAAGCATATTTAGAGAGCATTTTGAACGTAAGTGTGTTGCCAATCCAGCATTT	1686		
Db	787	GACGAGTCCCAAGACTCTATTTTGGACAAAACGCGAAGATGAAGTGTGCGCATACCCA	846	QY	1687	TTGAAAGCTTCTGAAAAGCATATAGCATAGCCCTCCTCCTGAAAGATGCTTCCAGCTCCTCAGC	1746		
QY	847	CTGTACATCTTATGAAAAATGGAGAGACTTTTCTTCTGACATACATCAGTCTGTACACAGC	906	Db	1687	TTGAAAGCTTCTGAAAAGCATATAGCATAGCCCTCCTCCTGAAAGATGCTTCCAGCTCCTCAGC	1746		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 08:05:58 ; Search time 405 Seconds  
(without alignments)  
9607.570 Million cell updates/sec

Title: US-09-684-890A-1

Perfect score: 2378

Sequence: 1 gataaaccccaagacacaa.....tataaaagtgtttgatg 2378

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2341.6	98.5	2392	4	US-09-620-312D-261
2	2312	97.2	2343	4	US-09-949-016-4105
3	315.4	13.3	37565	4	US-09-949-016-15847
c 4	174	7.3	601	4	US-09-949-016-145422
c 5	174	7.3	601	4	US-09-949-016-145423
c 6	174	7.3	601	4	US-09-949-016-145424
c 7	168	7.1	8374	4	US-09-949-016-11911
c 8	167	7.0	601	4	US-09-949-016-145428
c 9	138	5.8	601	4	US-09-949-016-145421
c 10	105.6	4.4	601	4	US-09-949-016-145431
c 11	88	3.7	601	4	US-09-949-016-145425
c 12	87.4	3.7	601	4	US-09-949-016-145426
c 13	87	3.7	601	4	US-09-949-016-145430
c 14	76.4	3.2	601	4	US-09-949-016-145432
c 15	61.2	2.6	7218	1	US-08-232-463-14
c 16	45	1.9	1702	4	US-09-799-451-574
c 17	40.2	1.7	832	4	US-09-621-976-2813
c 18	39.8	1.7	1158	4	US-09-710-279-69
c 19	39.8	1.7	1353	4	US-09-710-279-1919
c 20	39.8	1.7	1566	3	US-09-134-801C-2802
c 21	39.8	1.7	3055	4	US-09-710-279-3748
c 22	39.8	1.7	3226	4	US-09-710-279-4320
c 23	39.8	1.7	3229	4	US-09-710-279-4234
c 24	38.2	1.6	832	4	US-09-621-976-2813
c 25	38.2	1.6	1141	4	US-09-806-708B-22
c 26	38.2	1.6	1689	4	US-09-902-540-6020
c 27	38.2	1.6	1722	4	US-09-902-540-232

Sequence 8976, Ap  
Sequence 8976, Ap  
Sequence 22, Appl  
Sequence 900, Appl  
Sequence 1, Appli  
Sequence 12822, A  
Sequence 1524, A  
Sequence 17074, A  
Sequence 17393, A  
Sequence 22174, A  
Sequence 13679, A  
Sequence 11944, A  
Sequence 15690, A  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 16291, A  
Sequence 8070, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-620-312D-261  
; Sequence 261, Application US/09620312D  
; Patent No. 6569662

##### GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 261  
; LENGTH: 2392  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (49)..(2274)  
US-09-620-312D-261

Query Match 98.5%; Score 2341.6; DB 4; Length 2392;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2353; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 7 CCCCAAGCACAAACATACCTTTTCGAGCAGTTGGCCCAAGATGGCGCCGCGGCGGA 66  
Db 7 CCCTCGAGGCCAAGATTTCGCACAGAGCAGTTCGCGCCCAAGATGGCGCCGCGGCGGA 66

QY	67	CCGCTGGCGACGCGAGCTGTGGCAGACCTGGCTTCTTAACACAGTCGTGTCTTGGCG	126
Db	67	CCGCTGGCGACGCGAGCTGTGGCAGACCTGGCTTCTTAACACAGTCGTGTCTTGGCG	126
QY	127	CTCCGGGAGGACGTGAAAAACGAGAGTCCAACCGAAGCTGAGAAAAACAGCTTCTCGTCG	186
Db	127	CTCCGGGAGGACGTGAAAAACGAGAGTCCAACCGAAGCTGAGAAAAACAGCTTCTCGTCG	186
QY	187	TTGCCCTTCGTCGCGCGCGCGAGTTGTGTACGAGAAAAAGTGTCTTTGGCCCTCGCGGGA	246
Db	187	TTGCCCTTCGTCGCGCGCGCGAGTTGTGTACGAGAAAAAGTGTCTTTGGCCCTCGCGGGA	246
QY	247	GAGCTTTTCTGTGGGACGAGAGACAGCTCTCTTCTTAGTCTGTCCCTTTCGGGGCCCC	306
Db	247	GAGCTTTTCTGTGGGACGAGAGACAGCTCTCTTCTTAGTCTGTCCCTTTCGGGGCCCC	306
QY	307	AGCGGCGGCGGGAAGAGCGCGCTGTGCCAGTACACAGAGATGTCTTGGCATAAATCCA	366
Db	307	AGCGGCGGCGGGAAGAGCGCGCTGTGCCAGTACACAGAGATGTCTTGGCATAAATCCA	366
QY	367	CCCTCTGTTGAAATCTATCAAGTCTTGTAAAGCCCAACACATCATGTAGCATTATA	426
Db	367	CCCTCTGTTGAAATCTATCAAGTCTTGTAAAGCCCAACACATCATGTAGCATTATA	426
QY	427	GGAAATAAAGGACTTATGTTAGTAAATTTACCTAAAGATGGGGGAAGAAATCTGAATTT	486
Db	427	GGAAATAAAGGACTTATGTTAGTAAATTTACCTAAAGATGGGGGAAGAAATCTGAATTT	486
QY	487	GAAAGTGGAAAAATCAACAGTGAATTTGTAGTACCACTCCAGTTGCGGAGAGATTTTTCACC	546
Db	487	GAAAGTGGAAAAATCAACAGTGAATTTGTAGTACCACCTCCAGTTTGGGAGAGATTTTTCACC	546
QY	547	AGTTCCACCTCTCTGACTCTTAAAGCATGCTGATGTATCCAGTGAATCTCTGGATCCC	606
Db	547	AGTTCCACCTCTCTGACTCTTAAAGCATGCTGATGTATCCAGTGAATCTCTGGATCCC	606
QY	607	CAGTGTGCTGTAAACATCAGACAAACGTAATCAGAAATTTACTCTCTAGTGGCGCGAG	666
Db	607	CAGTGTGCTGTAAACATCAGACAAACGTAATCAGAAATTTACTCACTAGTGGCGCGAG	666
QY	667	ACACCCACTAAACGTGATAATCTTTTTCAGAACCGGAAGAGAAAGTCTAGTCTCAATAAA	726
Db	667	ACACCCACTAAACGTGATAATCTTTTTCAGAACCGGAAGAGAAAGTCTAGTCTCAATAAA	726
QY	727	GGAAAGGCGTATACCCGATCTCTAGGAGAGACAGCAGTGTGCAATTTGGGCCATTG	786
Db	727	GGAAAGGCGTATACCCGATCTCTAGGAGAGACAGCAGTGTGCAATTTGGGCCATTG	786
QY	787	GACGCACTCCCAAGACTCTTATTGGACAAAAACGGCAAGATGAAGTAGTGGCATACCCA	846
Db	787	GCAGCAGTCCCAAGACTCTTATTGGACAAAAACGGCAAGATGAAGTAGTGGCATACCCA	846
QY	847	CTGTACATCTTATGAAATGGAGAGACTTTTCTGACATACATCAGTCTGTGTACACAGC	906
Db	847	CTGTACATCTTATGAAATGGAGAGACTTTTCTGACATACATCAGTCTGTGTACACAGC	906
QY	907	CCTGGAAATATTGGAAAGCTGTGGTCCATTGGCCCATGCTCCGCTGGAAGATAAC	966
Db	907	CCTGGAAATATTGGAAAGCTGTGGTCCATTGGCCCATGCTCCGCTGGAAGATAAC	966
QY	967	TATGTTTATGATCGTGTGTACTCTGTCTACCTGTGTCCCAATATCTTAGTGATC	1026
Db	967	TATGTTTATGATCGTGTGTACTCTGTCTACCTGTGTCCCAATATCTTAGTGATC	1026
QY	1027	GCTACTGAATCAGGAATGCTGTATCATCTGTGTGCTGTAGAGGGGAAGAGATGAC	1086
Db	1027	GCTACTGAATCAGGAATGCTGTATCATCTGTGTGCTGTAGAGGGGAAGAGATGAC	1086
QY	1087	CACAGCTCAGAAAGTCTCTGGGATTCAGGATGACCTCATCTCTCTGTATGTGTTT	1146
Db	1087	CACAGCTCAGAAAGTCTCTGGGATTCAGGATGACCTCATCTCTCTGTATGTGTTT	1146
QY	1147	GAAATGTTGAGTTGGAGCTTGCTTTTGAACCTGGCATCTGGAGGATGACCCCTTTTGAT	1206

Db	1147	GAAATGTTGAGTTGGAGCTTGCTTTTGAACCTGGCATCTGGAGGATGACCCCTTTTGAT	1206
QY	1207	TCTGACTTTTCTGTTCAGTCAAACTTTCATAGAGATCCCAAGTGTCTTCAAGATATCAC	1266
Db	1207	TCTGACTTTTCTGTTCAGTCAAACTTTCATAGAGATCCCAAGTGTCTTCAAGATATCAC	1266
QY	1267	TGTACTCATGAAGCTGGTGTACATAGTGTGGCTAACTTTGGATTTCATAAACTTTCACAA	1326
Db	1267	TGTACTCATGAAGCTGGTGTACATAGTGTGGCTAACTTTGGATTTCATAAACTTTCACAA	1326
QY	1327	TTTCTTGGATCAGTGAAGAAGATAGGATAGTTTACAGAACTCTCTACAGAACGAAA	1386
Db	1327	TTTCTTGGATCAGTGAAGAAGATAGGATAGTTTACAGAACTCTCTACAGAACGAAA	1386
QY	1387	TGCTTTGTTGAACACATCTTTTGTACGAGCCATTTGCCCTGCGAGGAGCTCCAAAT	1446
Db	1387	TGCTTTGTTGAACACATCTTTTGTACGAGCCATTTGCCCTGCGAGGAGCTCCAAAT	1446
QY	1447	CGAGGATTTGGATTGTACTGCACATTTCTGGGACCCACGATGATCTGCATCACAGTACC	1506
Db	1447	CGAGGATTTGGATTGTACTGCACATTTCTGGGACCCACGATGATCTGCATCACAGTACC	1506
QY	1507	TATGAATGCTCATATGCGCGTTATTAAGTACAGTCCATCCAGCGTCTCTCCCTGCTT	1566
Db	1507	TATGAATGCTCATATGCGCGTTATTAAGTACAGTCCATCCAGCGTCTCTCCCTGCTT	1566
QY	1567	TGTACTCGAGAGATGTTGAAGTGGCAGAGTCTTCCCTCGGTCTTGGCTGAAACCCCA	1626
Db	1567	TGTACTCGAGAGATGTTGAAGTGGCAGAGTCTTCCCTCGGTCTTGGCTGAAACCCCA	1626
QY	1627	GATTCCTTTGAAAAGCATATTAGNAGCATTTTGCACAGTAGTGTGCCAATCCAGCATTT	1686
Db	1627	GATTCCTTTGAAAAGCATATTAGNAGCATTTTGCACAGTAGTGTGCCAATCCAGCATTT	1686
QY	1687	TTGAAAGCTTCTGAAAAGGACATAGCCCCCTCTCTCTGAAAGATGCTTCAGCTCTCAGC	1746
Db	1687	TTGAAAGCTTCTGAAAAGGACATAGCCCCCTCTCTCTGAAAGATGCTTCAGCTCTCAGC	1746
QY	1747	AGAGCCACCCAGGTGTTCAGAGGAGTACATTTCTCAACAGGACTTTGCCAAGGAGGAG	1806
Db	1747	AGAGCCACCCAGGTGTTCAGAGGAGTACATTTCTCAACAGGACTTTGCCAAGGAGGAG	1806
QY	1807	ATTACGCGAGGCTCAAAATTTATGTGACCAAAAAAGAAACAACTAGAGATCTCAGT	1866
Db	1807	ATTACGCGAGGCTCAAAATTTATGTGACCAAAAAAGAAACAACTAGAGATCTCAGT	1866
QY	1867	TATFTGCGAGAGAGAGGAAAAAGTCTGCGGAAATGGCTGAGCGTTTAGCTGACAAATAT	1926
Db	1867	TATFTGCGAGAGAGAGGAAAAAGTCTGCGGAAATGGCTGAGCGTTTAGCTGACAAATAT	1926
QY	1927	GAGGAGCTTAAAGNAAAAACAAGGATATCATGACAGGATGAAAGAACTTCTCAGT	1986
Db	1927	GAGGAGCTTAAAGNAAAAACAAGGATATCATGACAGGATGAAAGAACTTCTCAGT	1986
QY	1987	TTTCACTCTGAGCTCCAGTTCTCTCTGATAGTACGAGACATGAAGAAAGATTTACAG	2046
Db	1987	TTTCACTCTGAGCTCCAGTTCTCTCTGATAGTACGAGACATGAAGAAAGATTTACAG	2046
QY	2047	CTGATACCTGATCAACTTCGACATTTGGGCAATGCCATCAACAGGTTTCTATGAAAAAG	2106
Db	2047	CTGATACCTGATCAACTTCGACATTTGGGCAATGCCATCAACAGGTTTCTATGAAAAAG	2106
QY	2107	GATTATCAACCAAGATGGAGAGGTTGAGTCTTCCAAAACCCACCATTTATCTC	2166
Db	2107	GATTATCAACCAAGATGGAGAGGTTGAGTCTTCCAAAACCCACCATTTATCTC	2166
QY	2167	AGTGCCCTACAGCAAAAGTGCATTCAGTCCATCTGAAAGAGAGGGTGAACATATAAGG	2226
Db	2167	AGTGCCCTACAGCAAAAGTGCATTCAGTCCATCTGAAAGAGAGGGTGAACATATAAGG	2226
QY	2227	GAAATGTTGAGTGGAGCTTGCTTTTGAACCTGGCATCTGGAGGATGACCCCTTTTGAT	2286



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 04:09:16 ; Search time 10297 Seconds  
(without alignments)  
11190.298 Million cell updates/sec

Title: US-09-684-890A-1  
Perfect score: 2378  
Sequence: 1 gataaacccacacagacacaa.....tataaaagggtgtttgatg 2378

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_ey.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2378	100.0	2378	9 HSNUP88	Y08612 Homo sapien
2	2341.6	98.5	2392	6 AR338770	AR338770 Sequence
3	2315.8	97.4	2360	6 CQ725479	CQ725479 Sequence
4	2301	96.8	2357	9 BC000335	BC000335 Homo sapi
5	2301	96.8	2366	6 AX828393	AX828393 Sequence
6	2267.8	95.4	2386	11 BV179863	BV179863 sqmml0705
7	1721	72.4	2418	10 BC032929	BC032929 Mus muscu
8	1703.8	71.6	2412	10 RN093692	U93692 Rattus norv
9	1700.4	71.5	2433	10 BC072524	BC072524 Rattus no
10	1698.8	71.4	2310	10 MMU532593	AJ532593 Mus muscu
11	945.4	39.8	2393	5 AJ617672	AJ617672 Xenopus 1
12	905.8	38.1	2735	5 BC077397	BC077397 Xenopus 1
13	904.2	38.0	2478	5 AJ617673	AJ617673 Xenopus 1
C 14	630.2	26.5	197577	2 AC124124	AC124124 Mus muscu
C 15	630.2	26.5	228883	10 AC127173	AC127173 Mus muscu
C 16	597.4	25.1	601	11 BV168085	BV168085 sqmml7240
17	523.8	22.0	746	4 AY553927	AY553927 Sus scrof
18	408	17.2	651	10 U01135	U01135 Mus musculu
C 19	315.4	13.3	118276	9 AC004148	AC004148 Homo sapi

C 20	217	9.1	250275	2	AC095695	AC095695 Rattus no
21	213.8	9.0	180274	2	AC027185	AC027185 Mus muscu
C 22	212.2	8.9	254116	10	ALS96136	ALS96136 Mouse DNA
23	200.6	8.4	201	11	BV198093	BV198093 sqmml9650
24	200.6	8.4	201	11	BV200393	BV200393 sqmml20405
25	200.6	8.4	201	11	BV200394	BV200394 sqmml20405
26	170	7.1	2344	3	AK112449	AK112449 Ciona int
27	167.6	7.0	560	11	BV178251	BV178251 sqmml99469
C 28	166	7.0	466	6	CQ070585	CQ070585 Sequence
C 29	166	7.0	466	6	CQ099041	CQ099041 Sequence
C 30	166	7.0	466	6	CQ137971	CQ137971 Sequence
C 31	166	7.0	466	6	CQ175148	CQ175148 Sequence
C 32	166	7.0	466	6	CQ221345	CQ221345 Sequence
C 33	166	7.0	466	6	CQ259446	CQ259446 Sequence
C 34	166	7.0	466	6	CQ297165	CQ297165 Sequence
C 35	166	7.0	466	6	CQ333529	CQ333529 Sequence
C 36	165.4	7.0	187	6	CQ079765	CQ079765 Sequence
C 37	165.4	7.0	187	6	CQ112093	CQ112093 Sequence
C 38	165.4	7.0	187	6	CQ150910	CQ150910 Sequence
C 39	165.4	7.0	187	6	CQ185006	CQ185006 Sequence
C 40	165.4	7.0	187	6	CQ234252	CQ234252 Sequence
C 41	165.4	7.0	187	6	CQ271955	CQ271955 Sequence
C 42	165.4	7.0	187	6	CQ309582	CQ309582 Sequence
C 43	165.4	7.0	187	6	CQ346195	CQ346195 Sequence
C 44	153	6.4	455	11	G27792	G27792 human STS S
C 45	149	6.3	546	6	AX780812	AX780812 Sequence

#### ALIGNMENTS

RESULT 1	HSNUP88	2378 bp	mRNA	linear	PRI 19-JUL-1999
LOCUS	Homo sapiens mRNA for nuclear pore complex protein 88 (Nup88 gene).				
DEFINITION	Homo sapiens mRNA for nuclear pore complex protein 88 (Nup88 gene).				
ACCESSION	Y08612.2	GI:5541878			
VERSION	nuclear pore complex protein; Nup88 gene.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Fornerod, M., van Deursen, J., van Baal, S., Reynolds, A., Davis, D., Murti, K.G., Franssen, J. and Grosveld, G.				
TITLE	The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88				
JOURNAL	EMBO J. 16 (4), 807-816 (1997)				
MEDLINE	97201523				
PUBMED	9049309				
REFERENCE	2				
AUTHORS	Fornerod, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-OCT-1996) M. Fornerod, St. Jude Children's Research Hospital, Department of Genetics, 332 N. Lauderdale, Memphis, TN 38105, USA				
REMARK	Revised by [3]				
REFERENCE	3 (bases 1 to 2378)				
AUTHORS	Fornerod, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-JUL-1999) M. Fornerod, St. Jude Children's Research Hospital, Department of Genetics, 332 N. Lauderdale, Memphis, TN 38105, USA				
COMMENT	On Jul 20, 1999 this sequence version replaced gi:1707521.				
FEATURES	Location/Qualifiers				
source	1. .2378				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="17"				
	/map="17p13"				
	/tissue_type="placenta"				
	/clone_lib="CloneTech Hu2002B#29203"				



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2347	98.7	2464	4	AAK52959		Human pol
2	2343.8	98.6	2464	4	AAK52960		Human pol
3	2341.6	98.5	2392	4	AAI58382		Human pol
4	2341.6	98.5	2392	5	ADQ98591		DNA encod
5	2341.6	98.5	2392	9	ADB48351		Novel hum
6	2336	98.2	2465	4	AAI60168		Human pol
7	2330.6	98.0	2393	4	AAK51976		Human pol
8	2330.6	98.0	2393	4	AAK51975		Human pol
9	2315	97.4	2346	13	ACN40104		Tumour-as
10	2301	96.8	2366	12	ADK60464		Angiogene
11	2301	96.8	2366	12	ADK60765		Angiogene
12	2301	96.8	2366	12	ADP73087		Angiogene
13	1710.2	71.9	2229	12	ADJ67702		Kifc DNA
14	1703.8	71.6	2412	10	ADB59152		Toxicity-
15	1703.8	71.6	2412	12	ADP72905		Renal tox
16	393.4	16.5	456	9	ACH28622		Human adu
17	322.2	13.5	429	9	ACH21235		Human adu
18	315.4	13.3	1084	13	ACN40103		Tumour-as
19	315.4	13.3	65854	4	AAK86282		Human imm
20	243.4	10.2	514	12	ACH69405		Human gen

